2012

The California Political Precinct Index: A Detailed Tool to Help Understand California Politics (Update for 2012)

David Latterman
University of San Francisco, dclatterman@usfca.edu

UNIVERSITY OF SAN FRANCISCO
Leo T. McCarthy Center for Public Service and the Common Good

The California Political Precinct Index: A detailed tool to help understand California politics

Update for 2012

David C. Latterman
Associate Director, Leo T. McCarthy Center for Public Service and the Common Good
dclatterman@usfca.edu

Contents

http://www.usfca.edu/centers/mccarthy/

Executive Summary

The California Political Precinct Index (CPPI) is a powerful political tool that can be used by both political practitioners and academics to better understand recent California voting trends. The CPPI has been updated for 2012 and significantly strengthened with a new analytical methodology. Modeled after the San Francisco Progressive Voter Index, the CPPI is a single-number index that ranks nearly all California voting precincts on a gradational 0-100 political scale, with 0 being the most politically 'conservative' precincts and 100 being the most politically 'liberal' precincts. The CPPI is constructed from California ballot measures that have ideological underpinnings, that are widely accepted by most of the professional and academic political community. For this update, ballot measures were taken from November 2008 and November 2010, and are strongly internally consistent — an improvement from the 2007 index.[1]

The 2012 CPPI covers virtually all of California, including all 58 counties and over 20,000 precincts. In this paper, I showcase this important tool, using it to examine selected California political trends, including:

- How California Assembly and Congressional District representatives line up with their district's CPPI
- An examination of political indicators and the 2010 Governor's race
- Observations of Latino voting trends
- The Latino voting bloc and selected ballot measures

I conclude with discussing some of the future examinations I and other USF researchers intend to perform. For instance, the CPPI will be a useful metric in predicting outcomes in California's new open primary system in the 2012 elections.[2]

[1] In this report, the variable is called CPPI_2011, since most of the work for this report was completed in late 2011.
[2] Special thanks to the McCarthy Center for their support of this project, and to Alia Al-Sharif for her research assistance.

Background

Single-number political indices can be extremely accurate in explaining and predicting voting behavior and political trends. The Cook Partisan Voter Index is a national example of this kind of index, using Presidential races to determine the partisan leanings of Congressional Districts.[3] The accuracy of these indices depends on using the correct inputs for the geography that is being analyzed. It is crucial to use *electoral* data from the actual geographies for which we are making an index; otherwise, an index is primarily guesswork.

The California Political Precinct Index (CPPI) is modeled after the San Francisco Progressive Voter Index (SFPVI), which is a 0-100 conservative-to-liberal scale of all San Francisco voting precincts.[4] This was originally created in 2002 by San Francisco State Professor Emeritus Rich DeLeon. There have been several iterations of the PVI, which is used by many San Francisco political workers to better target field and mail efforts, as well as to explain election results.[5,6]

The CPPI, like the SFPVI, uses ballot measure data at the precinct level for electoral inputs. The precinct is the smallest unit of data available for election results, and all voters in California are able to vote on statewide ballot measures. There are no county or local measures in the CPPI. In order to be able to assign an ideological meaning to the index (that is, defining what 'left' and 'right' mean), the measures and their results have to be interpretable within a framework that most political workers, and even lay people, understand. For instance, voters who support pro-choice legislation are considered more liberal than those who do not. Voters who oppose tax measures are generally thought to be more conservative.

It is important to note that candidate results do not go into the making of the CPPI, and therefore the index is more accurate in interpreting and predicting other ballot measures. Other attributes besides a candidate's political leanings can influence election results. For example, with candidates, personality matters as does the strength and skill of a campaign. However, this and previous iterations of the CPPI have shown that it can be very accurate in examining candidate results along its conservative-liberal lines.

Methodology

I present here a brief methodology for the CPPI. A more detailed explanation can be found in Appendix 1. The CPPI is a difficult undertaking because in most of California, electoral precincts change from year to year. Moreover, the precinct boundaries that are used for reporting the Statement of Vote often differ from those that define where someone is registered to vote. To create a proper index, all of the precincts need to be somehow standardized.

---

[3] http://cookpolitical.com/node/4201

[4] A bibliography of all the PVI reports since 2002 can be found at http://www.sfusualsuspects.com/resources/deleon_latterman/.

[5] 'The Powerful Progressive Voter Index', *Politics Magazine*, August 2009, 'Mission, St. Francis Wood Political Opposites', *San Francisco Chronicle*, 9/14/08

[6] 'Mission, St. Francis Wood Political Opposites', *San Francisco Chronicle*, 9/14/08

Taking data from UC Berkeley's Statewide Database (SWDB), I was able to collect precinct data from the November 2008 and November 2010 general elections. I chose these years because the SWDB has complete datasets for these time periods, and because they are recent enough that the new CPPI will be predictive for several years. They also have several viable ballot measures that are easily interpretable on a conservative-liberal continuum. This is not always the case for California ballot measures in a given year.

To standardize the precincts, I was able to break down each precinct from 2008 and 2010 into their constituent US Census blocks. The block was assigned the electoral data of the precinct in which it resided. The blocks were matched and reassembled into standardized base precincts — in this case, precincts from the November 2008 election. In past versions of the CPPI, I used a precinct construct from the SWDB called 'map precincts' to join precincts from year to year. The census block technique has proven to be more accurate and consistent. Furthermore, other data can be easily added to the dataset, including US census and voter registration data.

I carefully selected 9 ballot measures from 2008 and 2010, listed in Table 1. These were either social issues or fiscal measures that were easily interpretable. These measures had a very strong Cronbach alpha value of 0.92, a measure of internal reliability that tells us these measures are effectively measuring the same thing. The inter-item correlation was a solid 0.55.

Table 1: Propositions used in the 2011 CPPI

Proposition	Date	Description	Result
Prop 3	Nov 2008	$980 million in bonds for children's hospitals.	Fail
Prop 4	Nov 2008	Waiting period and parental notification before minors obtain abortions	Fail
Prop 5	Nov 2008	Treatment for youth; rehab and shorter parole for nonviolent adult drug offenders	Fail
Prop 8	Nov 2008	Eliminates the right of same-sex couples to marry	Pass
Prop 10	Nov 2008	$5 billion in bonds for alternative fuels	Fail
Prop 19	Nov 2010	Legalize and tax marijuana	Fail
Prop 21	Nov 2010	Increase vehicle license fees by $18 a year to fund state parks	Fail
Prop 23	Nov 2010	Suspend AB 32, the "Global Warming Solutions Act" until unemployment falls below 5.5% for a year	Fail
Prop 25	Nov 2010	Budget and related legislation can be passed with simple majority, rather than current 2/3^{rds} requirement	Pass

The index itself was created by a factor analysis with a varimax rotation. This analysis yielded two factor scores for each precinct, which were summed and then standardized to a 0-100 scale, where the most conservative precinct was assigned a score of '0' and the most liberal precinct was assigned a score of '100'. All of the precincts fit into this ideological continuum.

Because all of the data here are aggregated precinct data, we must be mindful of the ecological fallacy when conducting analyses with the CPPI. There are ways to convert aggregated data into individual

voter data, but for this paper and most other CPPI analyses I omit this.[7] The intent here is to show important voting trends, for which the aggregated data of the CPPI are suitable.

General results

The CPPI, though reaffirming what we understand to be standard California political trends, provides a very subtle view of the California body politic. Map 1 shows a statewide look at the 2011 CPPI. I use a convenient red-blue scale to illustrate the scale, but these aren't meant to necessarily infer party affiliation. As is consistent with California's recent political narrative, much of the coast is liberal, while inland precincts are generally more conservative; however, with precinct-level resolution the picture becomes more complicated. Large swaths of less-populated areas are indeed conservative, but areas of settlement are more liberal. For instance, along Route 99 in the Central valley, there is a string of 'blue' dots representing more liberal towns than the open farmland expanses around them.

[7] For example, see King, Gary (1997) A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data. Princeton: Princeton University Press

Map 1: The 2011 California CPPI

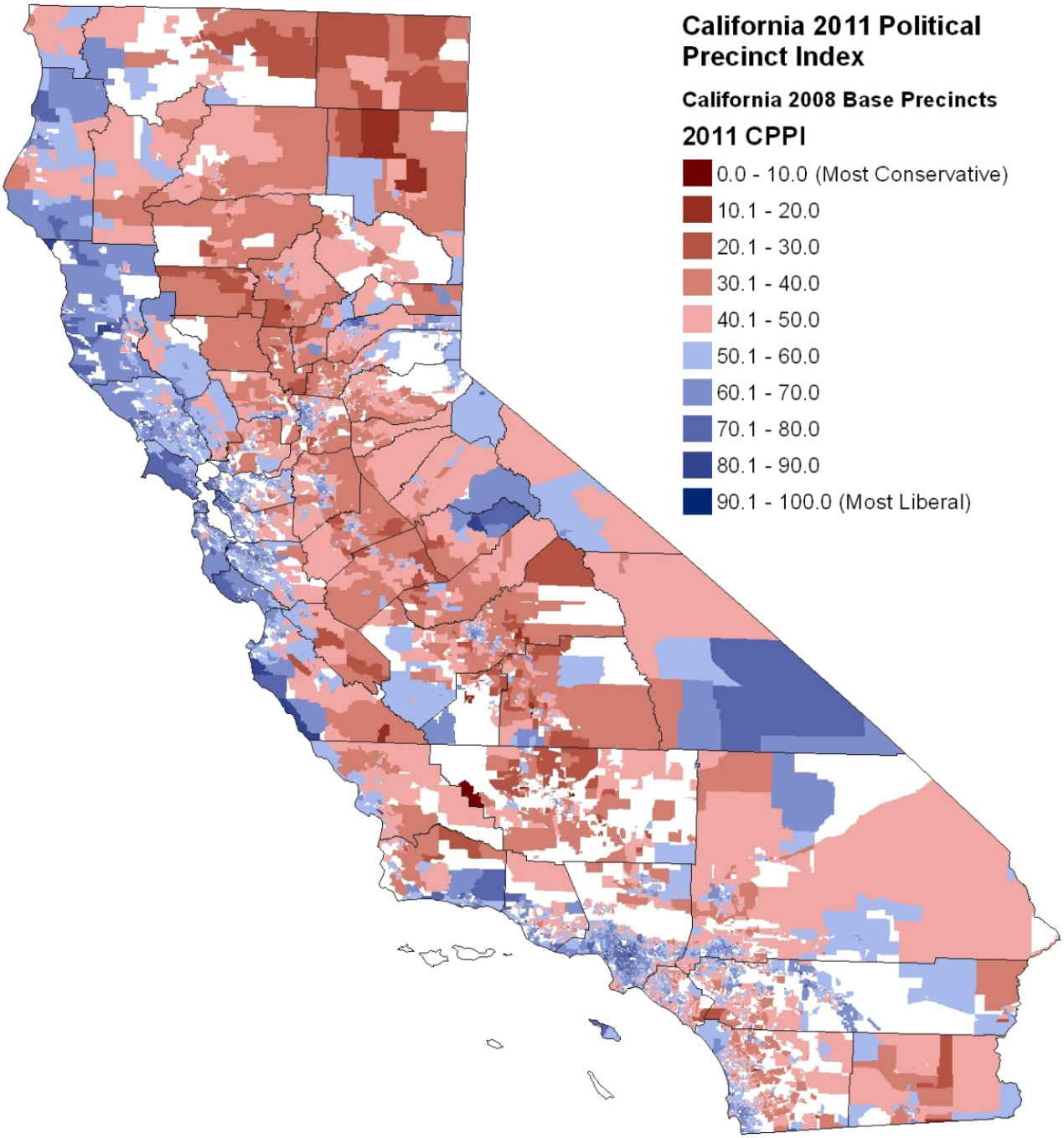
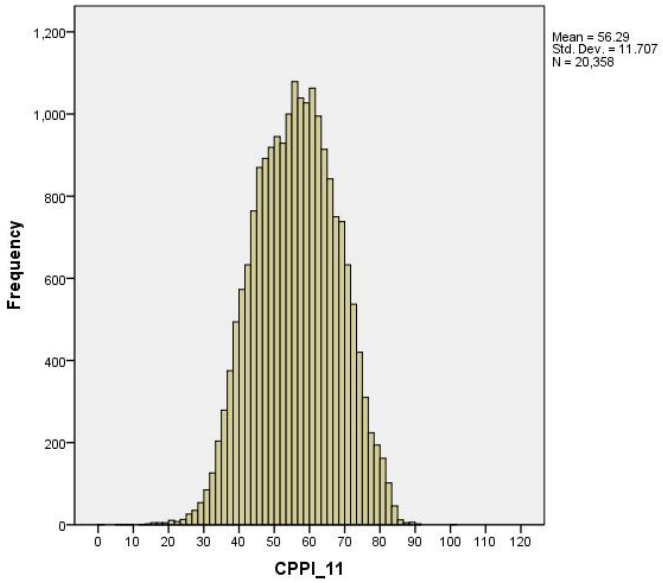


Figure 1 shows a histogram of the precinct results. The CPPI values of the precincts are normally distributed with a mean of 56.3. This is noteworthy because the mean is to the political left of CPPI=50, which is the midpoint of the ideological construct. Interestingly, more California precincts (and hence voters) are a little to the left of what could be called the ideological center of the state. This should not be surprising considering both the higher number of registered Democrats in the state, and the greater number of liberal precincts in the dense urban areas.

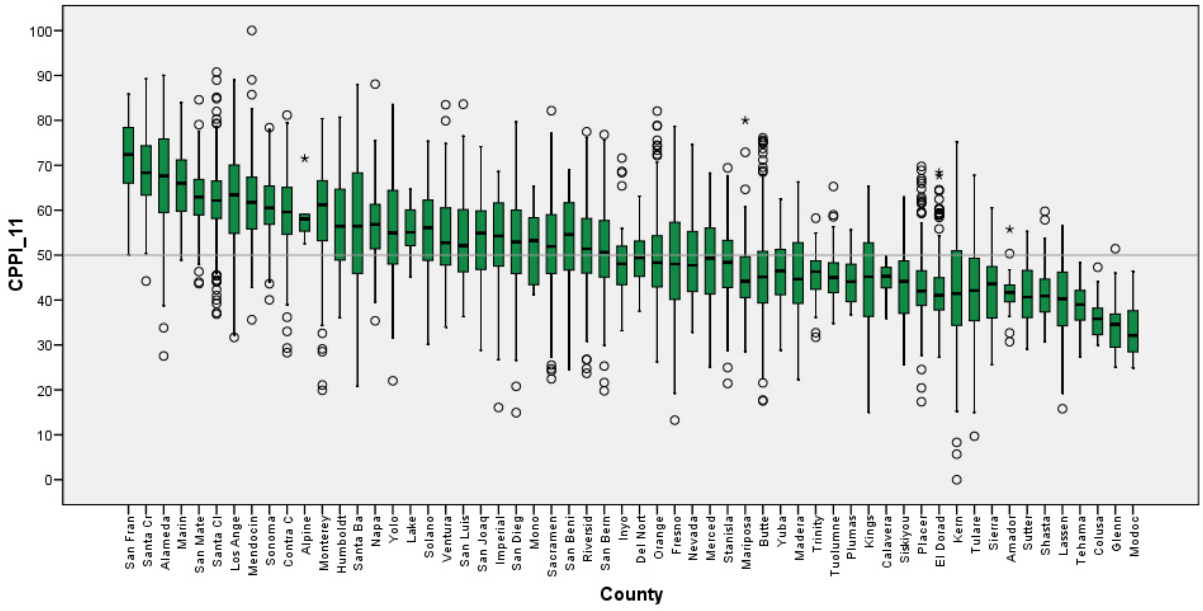
Figure 1: CPPI Histogram for all precincts



The first geographic aggregation that I examine with the CPPI is at the county level, which will be intuitive for many readers. Figure 2 displays a boxplot of county results, ordered from most liberal to most conservative. The Bay Area counties are the most liberal, while the counties of Northern California and the Central Valley are the most conservative. Counties considered 'right in the middle' include Riverside, San Benito, Inyo, and Del Norte.[8] A primer on how to read boxplots is in Appendix 2, and detailed data for the plots are available in Appendix 3.

[8] For a detailed examination of the politics of San Benito County, as most "centrist" county in California, see Cook, Corey and Latterman, David (2011) "San Benito County and California's Geopolitical Fault Lines," *California Journal of Politics and Policy*: Vol. 3: Iss. 1, Article 3.

Figure 2: Boxplot of county-aggregated CPPI values. Left to right is more liberal to more conservative



As elucidating as it is to look at the order of the county medians, the range of values within a county (or any aggregated geography examined here) is also noteworthy. For example, Santa Barbra and Kern Counties have large ranges, while other counties like Marin or Colusa have very narrow ranges. There are many reasons for this, like ethnic homogeneity, geography, and county industries. Ranges can be used to explain why some races are close while others or not, and understanding ranges can affect campaign strategy if a geography has many different voting blocs.

CPPI and candidate races

Even though the CPPI is constructed from ballot measure results, it is an accurate predictor of how precincts and regions will vote for candidates. Figures 3 and 4 display the correlation between CPPI and the percent vote for Governor Brown and Attorney General Harris in 2010, respectively.[9] There is also a geographic overlay on the correlation scatterplot.[10] The strong correlations ($R^2 > 0.8$) reveal the strength of the CPPI in explaining (and predicting) candidate results.

[9] The values for Brown and Harris use the standard *Two Party* Democrat-Republican percentages.
[10] Definition of regions:

Alpine	**Bay Area**	**C Coast**	**C Valley**		**LA Area**	**North**	**South**
Alpine	Alameda	Monterey	Amador	Merced	Los Angeles	Del Norte	Imperial
El Dorado	Contra Costa	San Benito	Butte	Sacramento	Orange	Humboldt	Riverside
Inyo	Marin	San Luis Obispo	Calaveras	San Joaquin	Ventura	Lassen	San Bernardino
Mono	Napa	Santa Barbara	Colusa	Stanislaus		Mendocino	San Diego
Nevada	San Francisco	Santa Cruz	Fresno	Sutter		Modoc	
Placer	San Mateo		Glenn	Tehama		Plumas	
Sierra	Santa Clara		Kern	Tulare		Shasta	
	Solano		Kings	Tuolumne		Siskiyou	
	Sonoma		Lake	Yolo		Trinity	
			Madera	Yuba			
			Mariposa				

The plots between CPPI and Governor, and between CPPI and Attorney General are quite similar, even though Brown did several points better than Harris. But, we see that Harris did a little more poorly in LA and the Central Valley at nearly all political levels, indicating precincts where statewide Republicans have a little more strength.

Figure 3: Correlation between CPPI and % Brown in 2010 with a geographic overlay

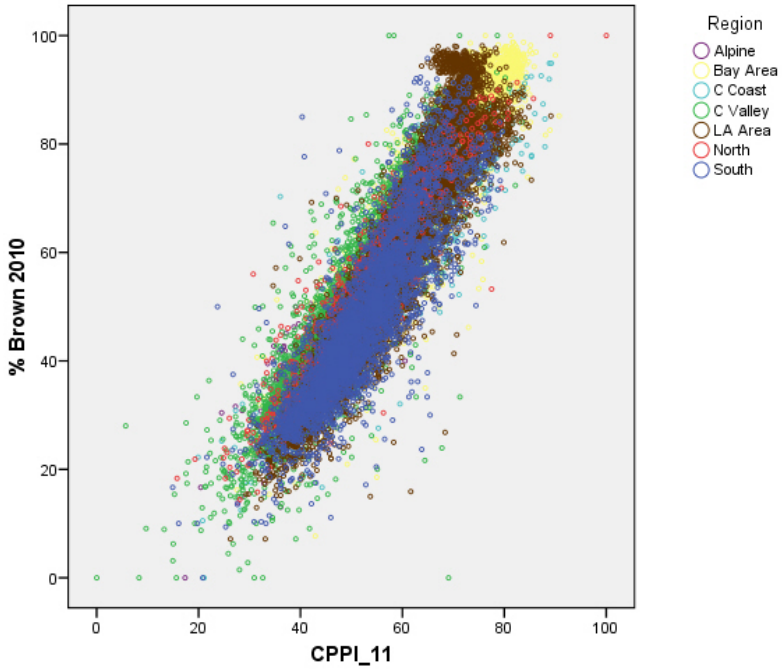


Figure 4: Correlation between CPPI and % Harris in 2010 with a geographic overlay

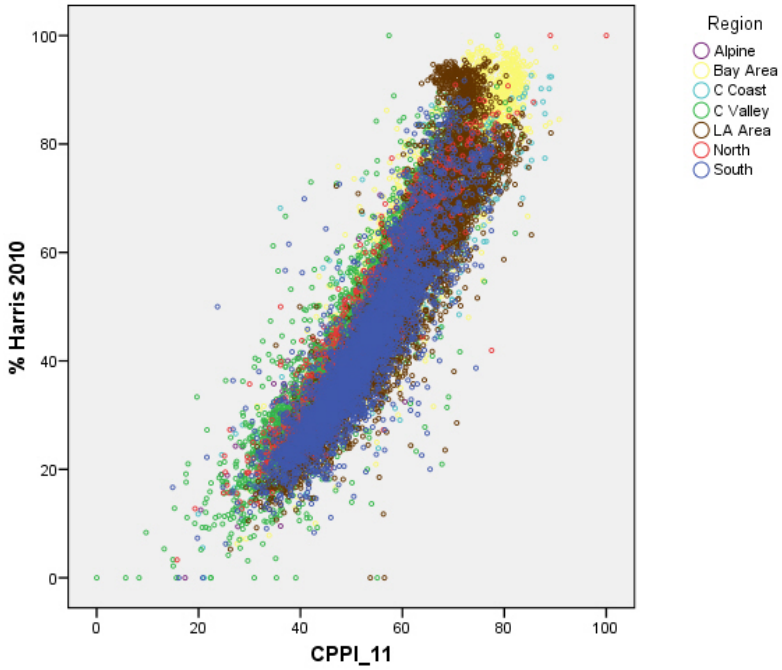
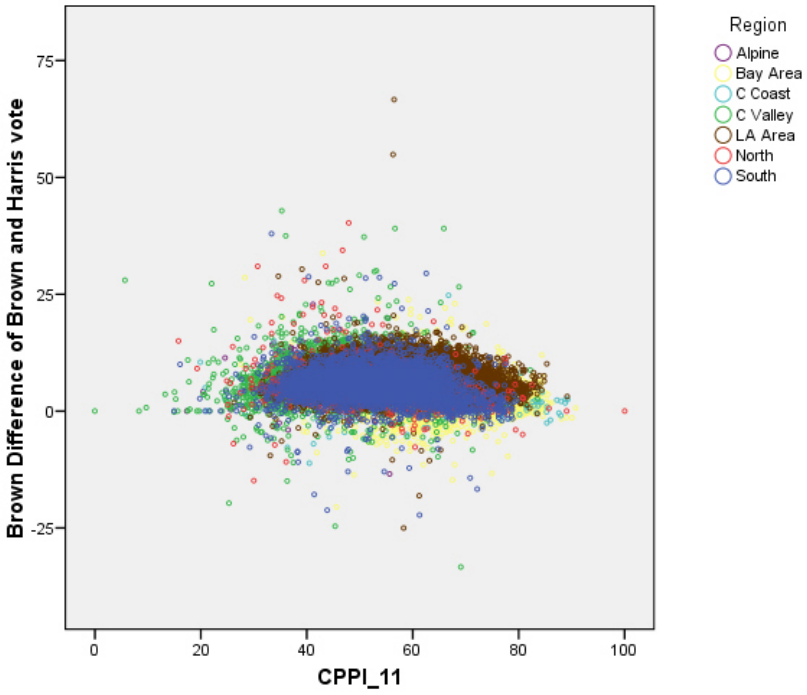


Figure 5 displays the correlation between CPPI and the *difference* between Brown's and Harris' percentage votes (Brown minus Harris). It is clear that the difference is nearly uniform across the political spectrum, with mild exceptions in the Central Valley and LA/Orange County region. Harris' Republican opponent was from LA, explaining her weaker performance there, and of course the Central Valley is quite conservative. Notwithstanding, the CPPI shows that politics in itself was not a strong factor in Harris' weaker showing throughout the state.

Figure 5: Correlation between CPPI and Brown minus Harris vote difference in 2010 with a geographic overlay



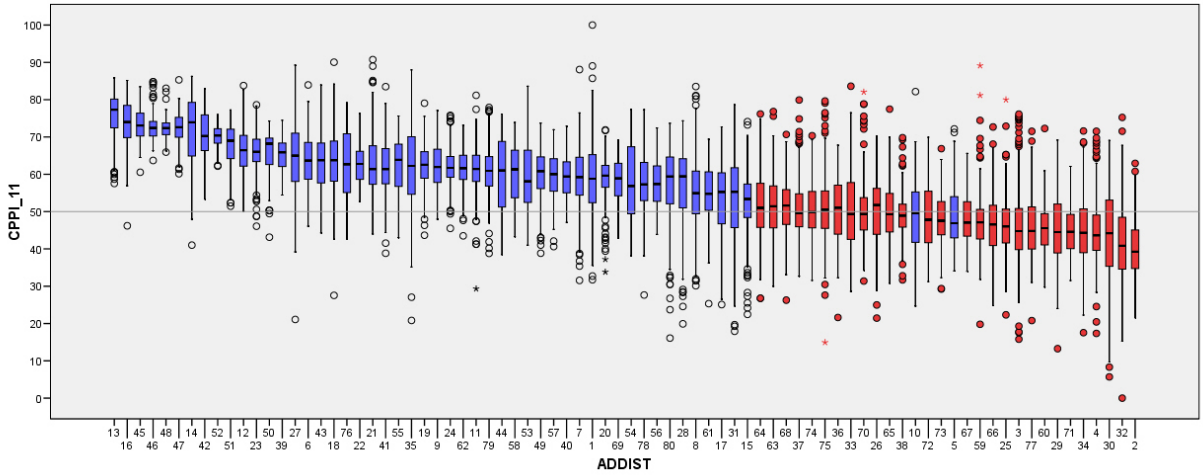
California State Assembly and Congressional Districts and the CPPI

Another way to use the CPPI is to examine the aggregate values of various political districts in order to see how the districts' elected representatives line up with the CPPI. If the CPPI is accurate, more liberal districts should have Democratic representatives, while more conservative districts should have Republican representatives. Moreover, the level to which districts are liberal or conservative should lend some insight into how easily partisan candidates win races.

Figure 6 shows a boxplot of the CPPI of California Assembly districts; the color of the boxes indicate whether the current representative is either a Democrat or a Republican. The graph shows a near-perfect alignment of high CPPI = Democratic Assemblymember and low CPPI = Republican Assemblymember, with the transition point right around CPPI = 50. The only two exceptions, where Democrats appear in the more 'conservative' portion of the graph, are AD 5 and AD 10, which were very

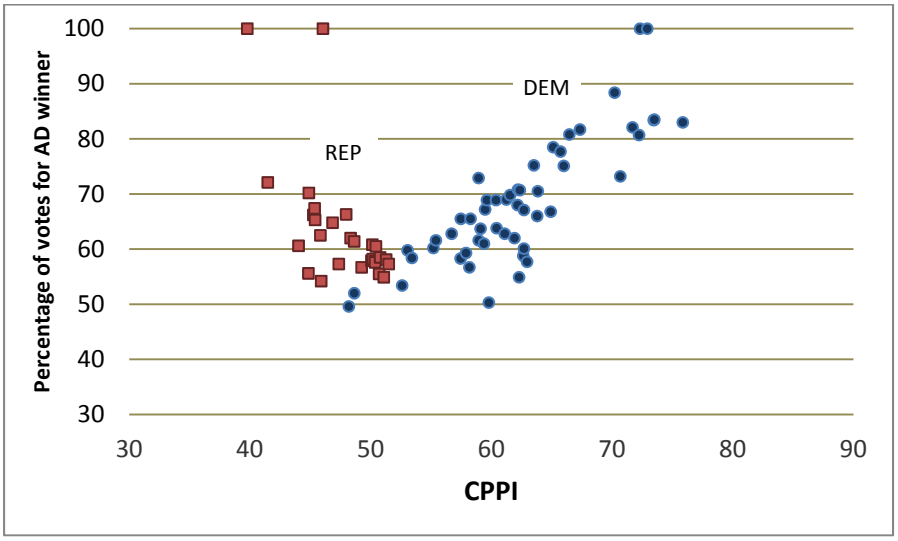
close races in the Sacramento area. The districts' CPPI were also still close to the midpoint, with AD5 having a mean CPPI = 48.2 and AD10 = 48.6. Appendix 4 shows the raw Assembly District data.

Figure 6: Boxplot of CPPI aggregated into California Assembly Districts, colored by representative party affiliation



A graph of CPPI vs. percent of the vote each Assemblymember last received shows the subtlety of the CPPI (Figure 7). The graph shows two distinct trendlines, originating from the CPPI midrange, indicating strong correlations between the strength of a district's ideology and how easily candidates win. The red markers represent Republican winners and the blue markers represent Democratic winners.11

Figure 7: Plot representing correlation between CPPI and the percent of vote that the Assemblymember received



[11] In both the California Assembly and Congressional plots, a few races saw uncontested races, resulting in receiving 100% of the ballot.

I perform the same analysis with Congressional districts (boxplot is Figure 8). The very consistent result is the same as with the Assembly, with a clear Democratic region, a transition area, and a Republican region of the graph. The transition area is right around CPPI = 53, indicating a slightly more liberal transition area than the Assembly. Appendix 4 shows the raw Congressional District data.

Figure 8: Boxplot of CPPI aggregated into California State Senate Districts, colored by representative party affiliation

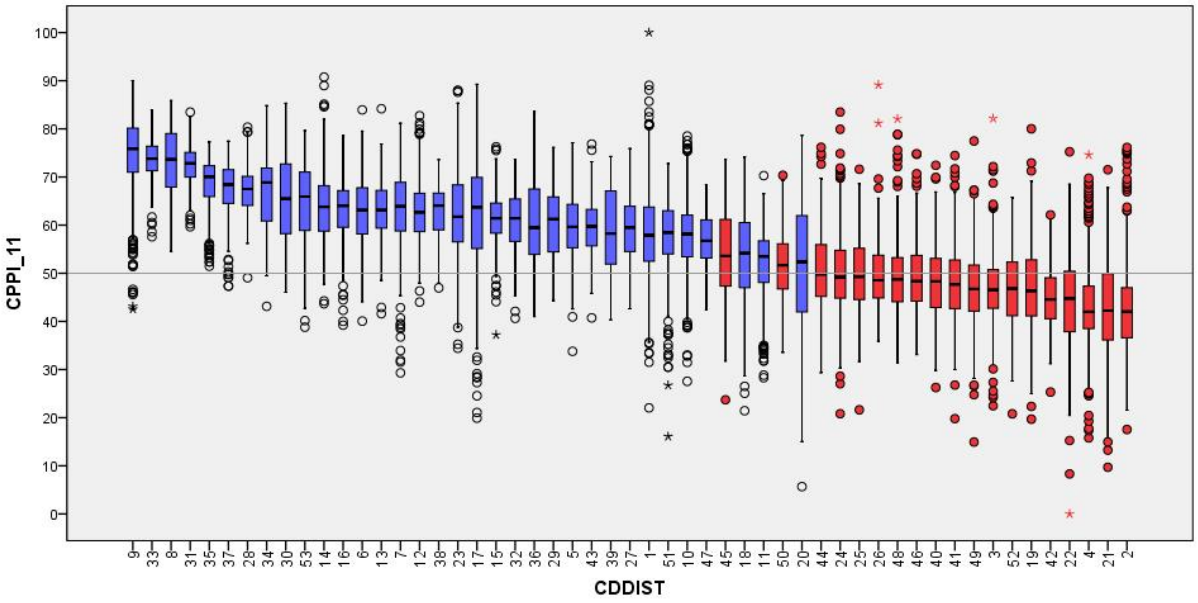
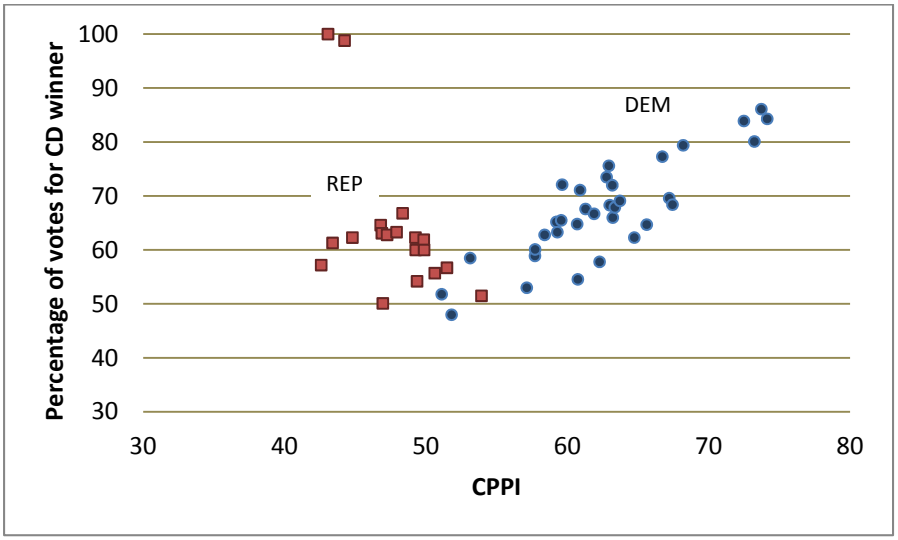


Figure 9: Plot representing correlation between CPPI and the percent of vote that the Congressperson received



CPPI and Latino voters

Much has been written on the Latino vote in California and its influence.[12] This demographic has often been considered a crucial and growing voting bloc. Based on the 2011 California voter file, around 24% of the overall electorate is Latino. Somewhat fewer of those voters could be considered frequent voters (19% of the November 2010 electorate was Latino), but Latino voters are still a sizable and increasing bloc.

Because of ecological fallacy issues, we are examining Latino trends by precincts. However, looking at precincts with high percentages of Latino voters lend some insight on overall Latino voting patterns. Table 2 shows the mean CPPI per percentage Latino categories. There is a clear upward trend that shows that the more Latinos in a precinct, the more liberal it is likely to be. Also, there is a strong trend that the higher the Latino percentage, the lower the standard deviation is, indicating that geographies with higher concentrations of Latino voters will have less deviation from a liberal voting platform.

Table 2: CPPI values for precincts of varying Latino percentages

Percent of Latino voters within precinct	Mean CPPI	Precinct N	Std. Deviation
0-25%	54.9	14925	12.2
26-50%	58.8	3459	9.5
51-75%	62.8	1156	6.7
76-100%	67.0	448	4.6

A more detailed geographic look at Latino voting trends show that Central Valley Latinos are generally more conservative than coastal and urban Latinos. (Table 3). CPPI values for high-Latino precincts are relatively similar, except for Central Valley precincts. The Central Valley CPPI mean values for high-Latino precincts are 6-8 points lower than other regions, perhaps reflecting some overprint of the more conservative political culture of the Central Valley. Figure 10 displays this graphically, for precincts greater than 40% Latino.

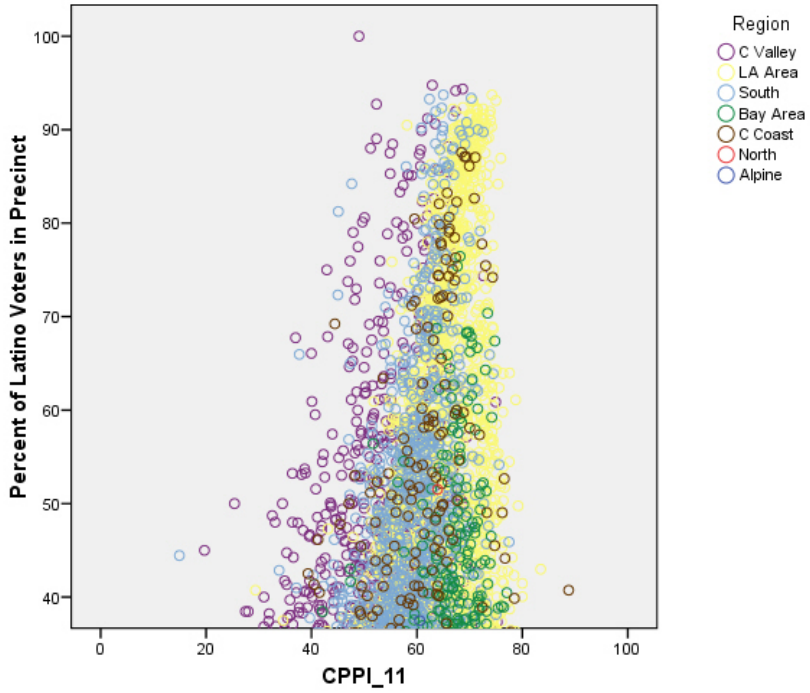
Table 3: Regional breakout of CPPI and percentage Latino per precinct

Percent of Latino voters within precinct	Region						
	Alpine	Bay Area	C Coast	C Valley	LA Area	North	South
	Mean	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*	*Mean*
0-25%	44.5	63.4	58.4	46.6	55.1	51.5	49.9
26-50%	59.8	66.7	57.4	52.0	61.8	63.8	56.2

[12] For some examples, see the work of Prof. Matt Barreto at the University of Washington, Melissa Michaelson at Cal State East Bay, and Mark Baldassare at the Public Policy institute of California. The late Andre Pineta did much strong work from the practitioner side of Latino research throughout California.

51-75%	.	67.9	62.4	56.2	65.3	63.9	60.8
76-100%	.	68.2	67.0	59.2	68.6	.	64.6

Figure 10: Correlation between CPPI and percentage Latino per precinct, for precincts with Latino percentage greater than 40



The Latino voting bloc and ballot measures

Looking at Latino voting trends a little more deeply, I use the CPPI to examine two controversial ballot measures in the past couple cycles: Prop 8 - Same-Sex Marriage, and Prop 19 - Marijuana Legalization. Prop 8 in 2008 was a ban on same-sex marriage, while Prop 19 in November 2010 was the legalization of marijuana in California.

Both of these measures were seen as important 'social values' referendums. Same-sex marriage has been fought at the state level for several years and is one of the most controversial topics in American politics. Meanwhile, California has been at the forefront of the marijuana legalization movement for many years. Despite California's 'liberal' reputation, the same-sex marriage ban was passed and the marijuana measure was defeated.

The CPPI is an ideal tool to look at these measures in detail. Figures 11 and 12 show the correlation between the CPPI and the 'no' vote on Prop 8, and the 'yes' vote on Prop 19, respectively. I choose to use the 'no' vote on Prop 8 to keep the correlation with CPPI in the same (positive) direction, where higher numbers on the X axis represent more liberal precincts. Both measures correlate relatively strongly with CPPI. The R^2 for the Prop 8 plot is 0.42 and the R^2 for the Prop 19 plot is 0.49.

Figure 11: Correlation between CPPI and percent yes on prop 8

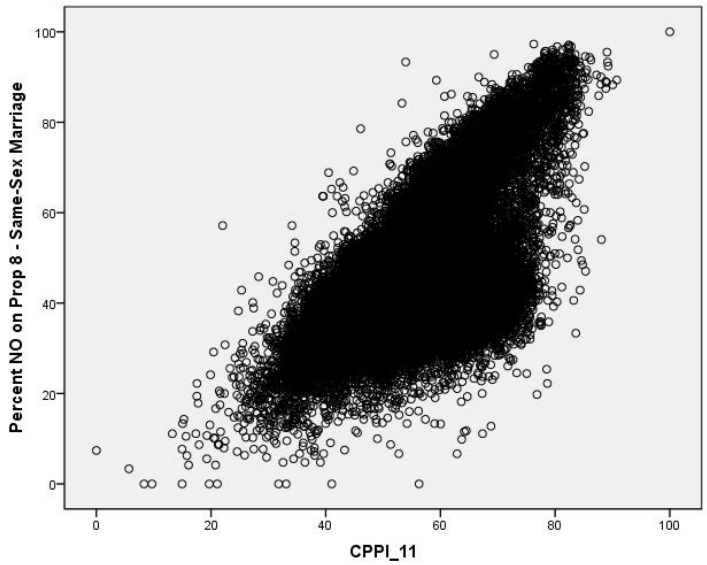
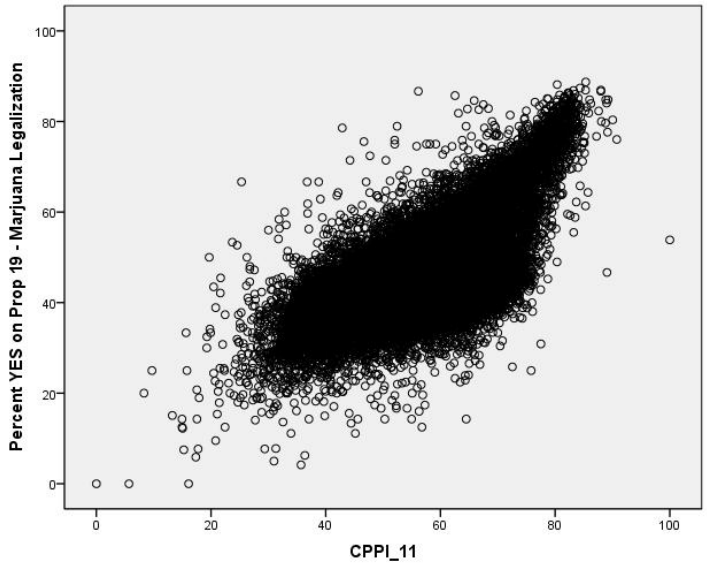


Figure 12: Correlation between CPPI and percent yes on prop 19



The most most notable feature about the two plots, and what brings the R^2 value down a little, is the fourth quadrant bulge in both plots, located in the same place. For both measures, this bluge indiates an otherwise liberal voting faction (higher CPPI values) that voted more conservatively on these measures, relative to what would have been expected given their CPPI values.

Figure 13: Correlation between CPPI and percent no on prop 8 with Latino categories

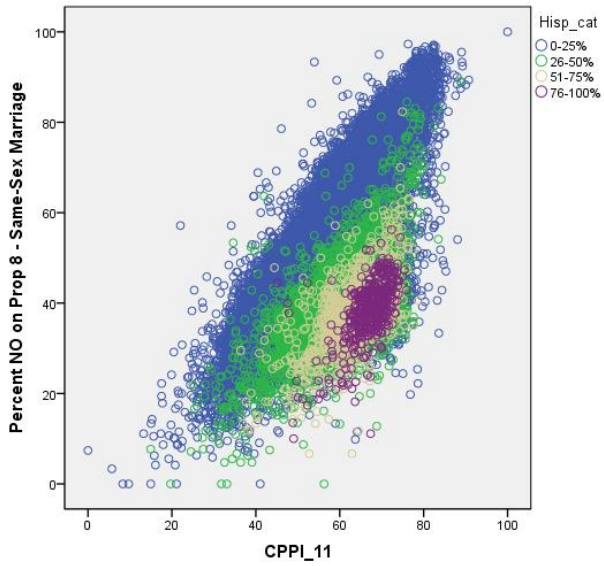
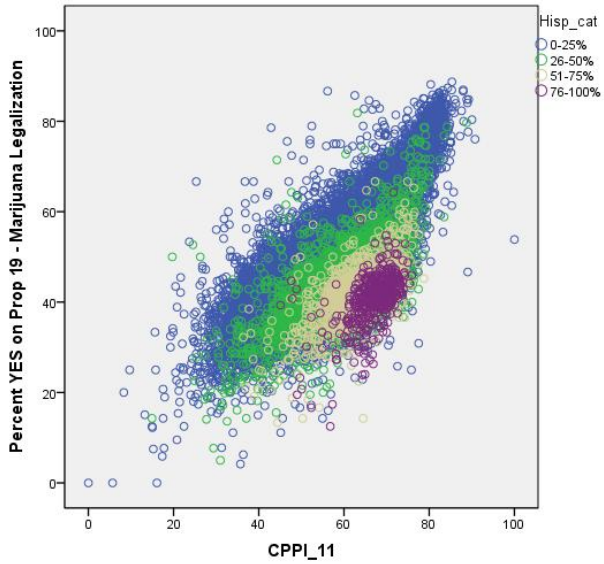


Figure 14: Correlation between CPPI and percent yes on prop 19 with Latino categories



The bulge in the plots correlates directly with the percentage of Latino voters in the precincts, indicating some degree of social conservatism within that voting bloc. Given that Latino voters are normally seen as solid Democratic, liberal voters, it should be clear that on some social issues, they — as a whole — do vote somewhat more conservatively. Table 4 displays this statistically, with the R^2 values of the correlations, and then the R^2 values controlled for the Latino percentage in the precinct. I also include a correlation to 2008's Prop 4 - parental notification of abortion - for another example. In all three cases, the correlative strength increases significantly once the correlations are controlled for the Latino population.

Table 4: R values for correlation between CPPI and selected measures, controlling the for percentage Latino portion of the precincts

	CPPI R^2	R^2 controlled for % Latino in precinct
Prop 4 Parental notification 2008	0.29	0.56
Prop 8 Same sex marriage 2008	0.42	0.64
Prop 19 Marijuana legalization 2010	0.49	0.64

Research questions and future work

The CPPI is a powerful tool that can be used to analyze and predict many types of California elections. It provides a subtle look at voting trends throughout the state, among the state's varied political geographies and their constituents. The CPPI can help interpret results on both candidate races and ballot measures, and can be aggregated into any California political geography and used for study.

This paper is meant to serve as an introduction to the CPPI. In the future, this index will be used at USF's McCarthy Center and other places in order to better understand California politics. Some of the research questions and future uses we wish to examine include:

- The effect of geography on political culture. How do geographic boundaries influence voting behavior? Do voters move to a place because it already has a political culture or do people impart their political values onto a place? We think these questions might be especially pertinent in specific urban neighborhoods.
- We intend to include the 2010 census into the CPPI dataset, which is comprised on the re-aggregation of census blocks. This will allow us a much deeper look at demographics and the CPPI. Some of this work was done successfully for the 2007 CPPI, using 2000 Census data.
- We intend to use John King's EZI software to attempt to link CPPI values with individual voters, or at least specific demographics.
- With the new 2012 open primary system for California elections, the CPPI will be an important tool in predicting the outcomes of the new primary system. In Spring 2012, we will publish a prediction guide of the Assembly, State Senate, and Congressional races, given the candidates in the race at that time.

About the Leo T. McCarthy Center at the University of San Francisco

The Leo T. McCarthy Center for Public Service and the Common Good is dedicated to inspiring and equipping students at the University of San Francisco (USF) to pursue lives and careers of ethical public service and service to others. The Center provides a non-partisan forum for education, service and research in public programs and policy-making. It supports undergraduate and graduate academic programs, including a Masters' Program in Public Affairs and an undergraduate Minor in Public Service. Through curricula that blends rigorous intellectual training with fieldwork experience, the McCarthy Center preparing students to articulate and promote the common good of all society's members through careers or service in government, non-profits or the private sector. Additionally, the McCarthy Center provides community-based learning opportunities both domestically and abroad, including the facilitation of government experiences for students.

The McCarthy Center values civic engagement and seeks to promote public interest research that encourages civil discourse and constructive interaction among the great diversity of residents and officials in the Bay Area. The Center strives to accomplish its goals by being transparent, nonpartisan and rigorous in designing its work and products.

Established in 1855, USF is San Francisco's oldest university. Consistently ranked as one of the most ethnically diverse universities in the country, the USF is committed to being a premier Jesuit Catholic, urban university with a global perspective that educates leaders who will fashion a more humane and just world. With nearly 9,600 students enrolled both on- and off-campus, the university offers undergraduate, graduate, and professional students the knowledge and skills needed to succeed as citizens and professionals, and the values and sensitivity necessary to be men and women for others.

CONTACTS FOR MORE INFORMATION:

For more information about the McCarthy Center, please contact Corey Cook, Director of the McCarthy Center, 415-422-6163, cdcook2@usfca.edu.

For more information about our public opinion research and research program designed to facilitate public discourse and civic engagement, please contact David Latterman, Associate Director of the McCarthy Center, 415-422-5960, dclatterman@usfca.edu.

Appendix 1: Detailed Methodology

The CPPI is a difficult undertaking because from year to year, precincts change in most counties in California. Moreover, there are often differences in Statement of Vote (SOV) precincts, which report election results, and registration precincts, which aggregate voter file data from registered voters. These all need to be consolidated into one standardized file, so all registration and SOV data are placed into consistent precincts.

Most of the data taken for this project are courtesy of the California Statewide Database (SWDB), hosted at UC Berkeley. They provide registration and SOV data for each California County for most elections - including primary, general, and special elections. More importantly, they often provided codebooks in order to standardize the precinct data into precincts from a selected year. Although using the codebooks to consolidate precinct data is a tedious process with potential error, it is likely still the most accurate method to undertake this kind of consolidation.

In previous iterations of the CPPI, I consolidated the precincts into what the SWDB calls 'Map Precincts', an artificial precinct construct that allows much of the precinct data to be placed into mapable polygons.[13] As long as there is a codebook for SOV or registration precinct-to-map precinct, the consolidation could occur. This process is relatively accurate, but occasionally leaves gaps in county data or specific election data.

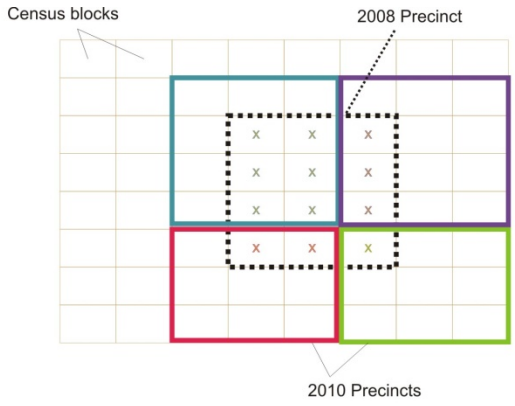
For the 2011 version of the CPPI, I utilized a complicated but more accurate consolidation system, where I broke all of the precincts into their constituent US Census blocks[14], assigned the precinct data to the blocks, and reassembled the blocks into a chosen base precinct, in this case the registration precincts from November 2008. Because this was a presidential election with a very high turnout, November 2008 provided a large number of blocks into which I could aggregate data from other elections. The SWDB provided block-to-precinct codebooks for all of the elections I used. Essentially, this procedure was a large database joining project, where Census blocks were the primary joining variable. Also, by using Census blocks, I could roll in selected Census data to provide an additional demographic component to the data.

The Census block technique worked especially well because the SWDB also provided a 'percent block' variable, which revealed how much of a block was inside a particular precinct. Usually, it was 100%, but not always. Since one block could only be assigned to one precinct, if the block wasn't 100% in one precinct, it was assigned the precinct where the largest portion of the block sat. In the rare cases where a block was split among many precincts, it was omitted from the dataset. I was then able to create a weighted average for a precinct based on how many people voted or were registered in a specific block. In other words, blocks with a lot of people counted more in compiling precinct data than a sparsely populated block.

[13] Usually, these precincts mimic the shapes for each county's registration precincts.
[14] Year 2000 Census

Precincts from November 2008 and November 2010 were broken into their constituent blocks, and reassembled into the resultant dataset. Technically, 2010 precincts were re-aggregated into 2008 precincts, since that had been chosen to be the base year. To this new dataset, I was able to attach census data, registration data, and SOV data. Any year's precincts can be deconstructed with the same technique, so the CPPI can be applied across many elections. See Figure 15 for an illustration.

Figure 15: Generalized diagram of how precincts from year-to-year were matched



1. The 2008 precinct is the 'base precinct' for the dataset
2. Census blocks marked with an x represent the blocks in this precinct
3. For the 2010 precincts, their respective data are attached all of the 2008 'x' blocks
4. If a block cuts across precincts lines, the precinct with the largest share is assigned to the block
5. The 'x' blocks, with 2010 data, are reaggregated to match the 2008 precinct.
6. This 2010 reaggregation is what is used to match the 2008 precinct data, creating a consistent dataset,

Once the dataset was assembled I carefully chose the ballot measures that would comprise the CPPI. I had 21 measures to choose from in 11/08 and 11/10. The chosen measures needed to have a general left-right divide, as understood by the political community and even lay people. There is a lot of discourse on what left-right means in California (or any political geography), but the issues had to be easily understood on this scale. For abortion, same-sex marriage, and other social issues, the divides are well understood, as it is for most bond measures. This makes no claims as to what people vote on these measures; after all, that's what the index is for. We solely need to interpret the measure itself into some basic left-right dichotomy.

Nine issues were ultimately chosen, which are listed in Table 1. All are relatively easily definable on the left-right scale.[15] The Cronbach's alpha, a reliability measure, for the nine issues was a very strong 0.920, with an inter-item correlation mean of 0.553. I then performed a factor analysis with varimax rotation on the issues, which yield two factors. I make no attempt here to interpret the relative loadings on the factors with their constituent issues. The factors were summed and the result was scaled from 0-100, with 0 being the more 'conservative' index value and 100 being the most 'liberal' value. Filtering

[15] Only precincts with more than 15 registered voters in both 2008 and 2010 were used for the actual computation of the CPPI.

the precincts for more than 15 registered voters introduced at least some level of statistical significance in the electoral results.

The index was back-correlated with the constituent measures for consistency. Results are in Table 5, which were quite strong, with no r value less than 0.529.

Table 5: Pearson correlation of the CPPI with its constituent measures. The 'percent no' values were used for this table.

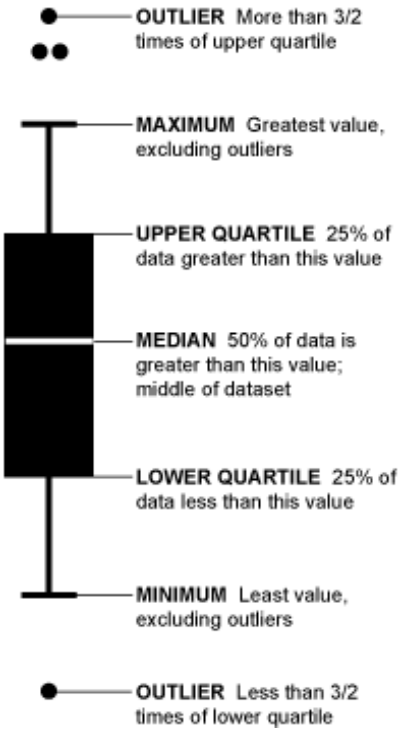
Measure	r value
p3n	-0.782
p4n	0.534
p5n	-0.884
p8n	0.645
p10n	-0.610
p21n	-0.701
p23n	-0.776
p25n	0.886
p27n	-0.904

It should be pointed out that even though I chose to use November 2008 as the base year, in truth any election could have been chosen, including November 2010. Mapping files were more easily available for November 2008, but as more data become available for November 2010, I may reconfigure the CPPI to set November 2010 as the base year.

Appendix 2: A note about boxplots

Boxplots are a convenient way to show aggregate summary data for a given category. In this case, boxplots show summary CPPI data for various geographies. The data displayed include the median, interquartile range, the full range of 'reasonable' data, and outliers. See Figure 16.

Figure 16: Diagram of a boxplot. Taken from http://flowingdata.com/2008/02/15/how-to-read-and-use-a-box-and-whisker-plot/



In this paper, the boxplots are ordered by highest to lowest median CPPI for the given geographies. Thus, the most 'liberal' geographies are on the left of the diagram.

Appendix 3: County-level aggregated CPPI data

Name	N	Mean	Median	Minimum	Maximum	Range	Std. Deviation
Alameda	909	67.3239	67.6361	27.55	90.04	62.49	10.17845
Alpine	5	59.3171	58.0490	52.56	71.52	18.96	7.28179
Amador	30	41.7813	41.6748	30.72	55.74	25.02	4.61257
Butte	138	46.1523	45.1632	17.54	76.13	58.58	11.79751
Calaveras	16	44.5766	45.2853	35.93	49.74	13.81	3.77065
Colusa	15	36.0390	35.8377	29.99	47.31	17.32	4.97881
Contra Costa	734	60.0195	59.6070	28.29	81.17	52.88	8.03295
Del Norte	16	49.2332	49.4012	37.55	63.11	25.57	6.32240
El Dorado	134	43.2372	41.0653	27.25	68.41	41.16	8.30550
Fresno	603	48.7090	48.0060	13.25	78.65	65.39	11.14719
Glenn	27	34.1243	34.5822	25.07	51.42	26.36	6.29798
Humboldt	126	57.3482	56.4225	36.07	80.71	44.64	9.91877
Imperial	100	53.3288	54.2661	16.09	68.61	52.52	10.37531
Inyo	26	49.3697	48.0580	33.17	71.60	38.43	9.87627
Kern	177	42.6681	41.4309	.00	75.24	75.24	12.91830
Kings	97	44.6210	45.1757	15.04	65.27	50.23	11.24260
Lake	40	55.4794	55.0872	45.08	64.65	19.56	5.08013
Lassen	35	40.1768	40.2825	15.79	56.52	40.73	9.76481
Los Angeles	4392	62.2521	63.4218	31.69	89.14	57.45	9.77541
Madera	65	45.6516	44.6601	22.33	66.30	43.97	10.11833
Marin	179	65.7189	65.9917	48.91	83.95	35.04	7.22459
Mariposa	21	47.3836	44.1951	28.56	80.01	51.45	12.63688
Mendocino	193	61.9328	61.7627	35.61	100.00	64.39	9.46569
Merced	112	48.2170	49.3336	25.08	68.27	43.19	9.44289
Modoc	19	33.0948	32.0984	24.89	46.43	21.54	6.73028
Mono	13	52.7620	53.2609	41.22	65.26	24.04	8.05343
Monterey	173	59.2873	61.2170	19.91	80.40	60.49	10.86191
Napa	140	56.7716	56.8499	35.38	88.07	52.69	7.26942
Nevada	82	48.2942	47.7324	32.85	74.59	41.74	8.13668
Orange	1787	48.8840	48.3257	26.26	82.07	55.81	7.91179
Placer	334	43.2891	42.0159	17.35	69.75	52.40	7.17087
Plumas	23	44.7731	44.0924	36.76	55.60	18.85	5.92680
Riverside	961	52.2308	51.4022	23.71	77.47	53.77	8.72190
Sacramento	652	52.6027	51.9577	22.46	82.16	59.70	9.38192
San Benito	52	52.5176	54.6059	24.55	69.02	44.46	10.53757
San Bernardino	1024	51.3739	50.6313	19.79	76.84	57.05	8.47146
San Diego	1873	53.1968	52.9298	14.94	79.64	64.69	10.26151
San Francisco	577	71.8478	72.3890	50.12	85.87	35.75	7.62107
San Joaquin	365	53.3765	54.9269	28.83	74.12	45.29	9.28001
San Luis Obispo	147	53.6600	52.1189	36.33	83.61	47.29	9.73141
San Mateo	523	62.7269	62.9401	43.67	84.58	40.91	5.91193
Santa Barbara	272	56.8220	56.4172	20.81	87.98	67.17	14.31270
Santa Clara	1062	62.3102	62.1699	36.85	90.75	53.91	6.47369
Santa Cruz	154	68.6836	68.3325	44.24	89.25	45.02	8.17755
Shasta	123	41.0236	40.9285	30.76	59.76	29.01	5.46442

Sierra	21	42.2339	43.6021	25.63	60.52	34.89	8.60035
Siskiyou	50	43.6605	44.1802	25.60	62.94	37.34	9.20925
Solano	177	55.2000	56.1245	30.13	75.36	45.23	9.87210
Sonoma	406	61.0700	60.5397	40.06	78.39	38.33	6.51939
Stanislaus	236	47.7751	48.4116	21.44	69.44	48.00	8.56911
Sutter	60	41.4897	40.6522	29.07	55.29	26.22	6.82753
Tehama	42	38.7974	38.9919	27.28	48.31	21.03	5.50218
Trinity	20	45.5474	46.3110	31.73	58.23	26.51	6.74319
Tulare	201	42.5735	42.1174	9.70	67.78	58.09	10.28592
Tuolumne	67	45.4940	45.0075	34.82	65.27	30.46	5.79042
Ventura	352	54.0897	52.7640	33.97	83.49	49.51	8.35630
Yolo	134	55.6898	54.9503	22.03	83.53	61.50	12.21525
Yuba	46	46.1038	46.5056	28.86	62.56	33.69	7.79943
Total	20358	56.2914	56.3951	.00	100.00	100.00	11.70706

Appendix 3: State Assembly and aggregated CPPI data

ADDIST	N	Mean	Median	Minimum	Maximum	Range	Std. Deviation
1	509	58.9431	58.8046	31.73	100.00	68.27	9.66259
2	372	39.7780	39.2405	21.56	62.94	41.38	7.33428
3	372	45.8945	44.7843	15.79	76.13	60.33	9.48934
4	423	44.8427	43.6371	17.35	71.52	54.17	7.82731
5	240	48.1877	46.8887	34.12	72.16	38.04	7.66255
6	325	63.8448	63.6869	46.09	83.95	37.86	6.68515
7	343	59.1028	59.2479	31.57	88.07	56.51	8.16781
8	234	55.4013	54.9401	30.13	83.53	53.40	9.89431
9	173	62.1667	61.9506	47.97	77.11	29.14	6.28618
10	287	48.6378	49.5762	24.69	82.16	57.47	8.52715
11	287	61.2366	61.4291	29.32	81.17	51.86	6.17701
12	282	66.4546	66.4646	50.12	83.79	33.67	6.17121
13	329	75.8539	77.3003	57.55	85.87	28.31	5.81260
14	336	71.6900	73.9557	41.01	86.24	45.24	9.08160
15	412	52.6039	53.3676	22.46	74.12	51.67	7.92637
16	262	73.4934	73.9954	46.20	85.16	38.96	6.23688
17	239	53.4145	55.2606	25.08	72.72	47.64	10.02371
18	253	63.5165	63.7735	27.55	90.04	62.49	7.76675
19	322	62.1978	62.5439	43.67	79.05	35.38	5.62509
20	223	58.9373	59.6350	33.82	71.89	38.06	6.05061
21	367	62.6932	61.4012	44.00	90.75	46.75	7.24538
22	222	62.6954	62.7872	52.65	76.27	23.62	4.80828
23	204	65.9938	66.0174	46.08	78.61	32.53	5.70677
24	283	61.9109	61.7288	45.50	75.82	30.32	5.00737
25	275	46.0388	46.0307	22.33	80.01	57.68	7.82499
26	249	50.1374	51.7486	21.44	70.28	48.84	9.25028
27	256	64.9111	64.9896	21.08	89.25	68.17	9.78635
28	234	56.7047	59.4335	19.91	74.35	54.44	10.10208
29	329	45.3538	44.4942	13.25	69.16	55.91	9.10673
30	219	44.0219	44.2233	5.69	69.06	63.38	12.78548
31	286	53.0503	55.3148	17.92	78.65	60.72	11.87392
32	111	41.4751	40.8194	.00	75.24	75.24	11.92029
33	253	50.3609	49.3475	28.61	83.61	55.00	10.18987
34	244	44.8743	44.3080	17.54	71.60	54.07	9.02802
35	250	62.2894	62.2486	20.81	87.98	67.17	11.26020
36	255	50.3649	51.0337	21.63	67.83	46.20	8.33668
37	235	50.7794	49.5643	32.69	79.91	47.22	7.50578
38	247	49.2481	48.9275	31.69	69.83	38.14	5.41152
39	118	65.1221	65.8836	54.45	74.48	20.03	4.19274
40	164	59.3807	59.3904	47.12	72.90	25.77	5.72729
41	257	62.6409	61.3912	38.82	83.49	44.67	7.18831
42	276	70.6763	70.2645	53.34	83.02	29.67	6.15915
43	195	63.7913	63.7622	44.30	83.92	39.63	8.51563
44	245	60.4280	61.0191	38.48	76.14	37.67	9.57463
45	148	73.4436	73.0669	60.54	83.47	22.93	4.34442

46	106	72.9128	72.4284	63.71	84.82	21.11	3.89971
47	215	72.2276	72.6261	60.16	85.31	25.15	4.19957
48	138	72.3119	72.3899	65.94	83.06	17.13	2.84601
49	173	59.6271	60.8152	38.84	73.70	34.87	6.59127
50	143	65.7225	68.2020	43.14	74.26	31.12	5.73166
51	193	67.3354	69.0189	51.46	77.27	25.81	6.08792
52	150	70.2047	70.4080	62.18	75.99	13.81	2.97233
53	264	59.7848	58.1111	41.00	83.59	42.59	9.40451
54	236	58.1746	56.8630	38.14	77.45	39.31	10.29079
55	209	62.3518	63.8762	43.02	75.62	32.60	7.62378
56	204	57.4630	57.3819	43.88	72.42	28.54	5.89226
57	177	59.4772	60.0305	40.62	71.94	31.31	6.00579
58	192	60.3813	61.3115	43.29	73.89	30.60	7.75657
59	293	47.3559	47.1135	19.79	89.14	69.35	7.49823
60	255	45.3940	45.5656	29.78	72.29	42.50	6.32462
61	167	55.1905	54.7859	25.32	69.40	44.08	7.41689
62	184	61.5451	61.6116	43.50	73.15	29.66	5.20549
63	283	51.2834	51.4214	29.97	76.84	46.87	7.37163
64	271	51.4924	50.9822	26.75	76.18	49.43	8.52762
65	322	50.0549	49.2658	30.80	77.47	46.67	8.01920
66	290	46.8412	46.5468	24.77	72.68	47.91	7.85878
67	300	47.9638	47.0438	34.00	65.61	31.61	6.20556
68	257	51.0861	51.6360	26.26	70.68	44.42	6.67315
69	144	58.2691	58.8958	42.80	69.27	26.47	6.16373
70	322	50.1662	49.3417	31.39	82.07	50.68	8.02239
71	281	45.2472	44.5595	31.53	62.12	30.59	6.45160
72	257	48.6313	47.8301	31.22	69.93	38.70	8.37395
73	247	48.3331	47.5452	29.22	66.87	37.65	6.42582
74	314	50.6971	49.8795	31.57	70.34	38.77	7.44110
75	317	50.4324	50.4648	14.94	79.64	64.69	8.52788
76	272	62.9521	62.6831	42.70	79.18	36.49	9.24031
77	318	45.8316	44.8165	20.80	71.53	50.73	8.02773
78	281	57.8962	57.2777	27.65	77.32	49.67	7.56318
79	187	61.1028	60.9431	38.79	77.92	39.13	7.03638
80	251	57.4467	59.4105	16.09	73.69	57.60	9.91553
Total	20358	56.2914	56.3951	.00	100.00	100.00	11.70706

Appendix 4: Congressional aggregated CPPI data

CDDIST	N	Mean	Median	Minimum	Maximum	Range	Std. Deviation
1	702	58.3882	57.9103	22.03	100.00	77.97	9.24776
2	526	42.5845	42.0267	17.54	76.13	58.58	8.85951
3	413	46.9445	46.5277	22.46	82.16	59.70	7.21217
4	680	43.3901	42.0139	15.79	74.59	58.79	8.02556
5	285	59.6253	59.6347	33.79	77.11	43.32	6.90462
6	511	63.2066	63.1460	40.06	83.95	43.89	6.95171
7	336	62.9933	63.9067	29.32	81.17	51.86	8.81404
8	471	73.2232	73.6751	54.53	85.87	31.34	6.81262
9	450	74.1294	75.8591	42.55	90.04	47.49	8.11117
10	495	57.7028	58.1448	27.55	78.51	50.96	7.55078
11	496	51.8067	53.4926	28.29	70.28	41.99	7.25710
12	470	62.9283	62.6581	44.00	82.73	38.73	6.33128
13	337	63.1737	63.1202	41.59	84.18	42.59	5.83832
14	467	63.7126	63.7729	43.67	90.75	47.08	6.90601
15	392	61.2822	61.4029	37.21	76.27	39.07	5.04566
16	381	63.3534	64.0311	39.22	78.61	39.39	6.64769
17	323	61.8843	63.7206	19.91	89.25	69.34	12.04046
18	332	53.1241	54.1712	21.44	74.12	52.68	10.05596
19	453	46.7916	46.3558	19.69	80.01	60.32	8.83953
20	314	51.1045	52.3586	5.69	78.65	72.96	13.02220
21	447	43.0666	42.2144	9.70	71.51	61.81	9.78487
22	237	44.2367	44.7506	.00	75.24	75.24	10.38236
23	333	62.2740	61.7475	34.42	87.98	53.56	9.82948
24	367	49.8646	49.1938	20.81	83.49	62.67	8.89833
25	428	49.8379	49.2691	21.63	71.60	49.98	7.48515
26	389	49.3675	48.5567	35.90	89.14	53.24	6.69230
27	279	59.2221	59.5100	42.62	75.91	33.29	6.51618
28	250	67.1994	67.5108	49.06	80.34	31.28	4.68448
29	308	60.6872	61.2381	44.30	76.14	31.85	7.54400
30	422	65.6077	65.5131	46.06	85.31	39.24	8.84610
31	201	72.4806	72.8548	59.65	83.47	23.82	4.31526
32	258	60.8955	61.4137	40.62	73.70	33.08	6.25756
33	288	73.7143	73.7947	57.57	83.92	26.35	4.32059
34	204	66.7134	68.8652	43.14	84.82	41.68	7.69781
35	268	68.1871	70.0351	51.46	77.27	25.81	5.73782
36	352	60.7367	59.4867	41.00	83.59	42.59	8.98908
37	293	67.4303	68.4332	47.32	77.45	30.14	5.63937
38	251	62.7696	64.0395	46.96	73.59	26.62	5.80966
39	285	59.2891	58.2470	40.34	74.26	33.92	8.44510
40	403	48.3453	48.3163	26.26	72.42	46.16	7.50705
41	499	47.9178	47.6895	19.79	74.45	54.66	7.87096
42	399	44.7916	44.5743	25.32	62.12	36.80	5.87944
43	293	59.5612	59.7397	40.74	76.84	36.10	5.89930
44	373	50.6171	49.8046	29.36	76.18	46.81	7.96134
45	398	53.9193	53.5693	23.71	73.69	49.98	9.21298

46	465	49.2491	48.3381	33.10	74.84	41.74	6.92374
47	236	57.1232	56.7593	42.39	68.36	25.97	5.47202
48	467	49.2711	48.7441	31.39	82.07	50.68	7.70075
49	434	47.2476	46.7393	14.94	77.47	62.53	8.01919
50	480	51.4761	51.6687	33.57	70.34	36.77	6.79381
51	372	57.7142	58.4446	16.09	72.82	56.73	7.87117
52	478	46.8777	46.7903	20.80	65.77	44.97	7.55535
53	367	64.7318	65.9383	38.79	79.64	40.84	8.62798
Total	20358	56.2914	56.3951	.00	100.00	100.00	11.70706